LA-UR-13-24646

Approved for public release; distribution is unlimited.

Title: Integration of Hyperspectral Imagery and Biosensors for Biological and Chemical Facility Classification

Author(s): Munsky, Brian
Yeager, John D.
Sellars, Scott
Mukundan, Harshini
Nadler, Brett Ross

Intended for: General Distribution

Issued: 2013-06-25

Integration of hyperspectral imagery and biosensors for biological and chemical facility classification

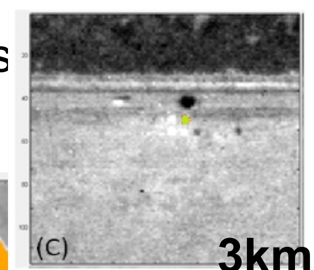
Brian Munsky, Scott Sellars, John Yeager

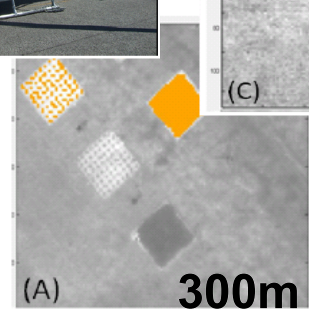
Mentors: Harshini Mukundan and Brett Nadler

4-25-13

Bio-Chem Facility Detection Project

"The goal of this project is to develop novel methods and identify novel signatures to indirectly determine whether or not these facilities are benign or are engaged in the manufacture of malicious chem/bio





Outline

- **Introduction**
- **Project Goals (proposed innovation)**
- **Background**
- **Methods**
  - Hyperspectral Imaging and Machine Learning
  - Detection Technology
  - Biosensor Development
- **Expected Results**
- **Project/Plan and Team**
- **Budget**

Introduction

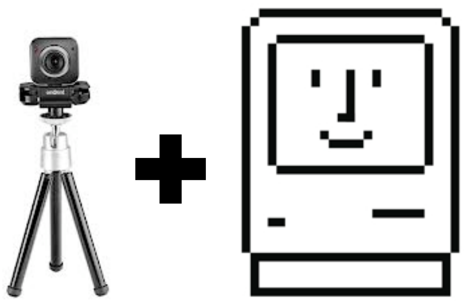
- When it comes time for a decision, there is no better intelligence, than human intelligence.
- How do you detect a biological and chemical manufacturing facility?
- Ideal scenario?
  - Constant observations and surveillance
  - Direct Measurements
  - Complete background of the people in the facility
- State of the art
  - Satellite and UAS platforms produce massive amounts of imagery from different wavelengths
  - Imagery signatures used in understanding explosives
  - Synthetic biological circuits can be de- signed

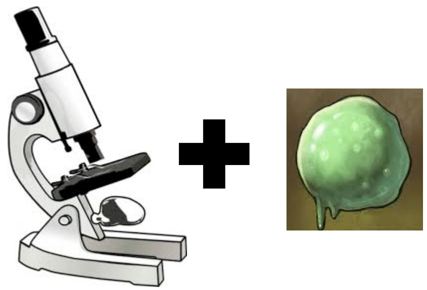

Google
Imagery ©2013 , Champaign County GIS Consortium, DigitalGlobe, USDA Farm Service Agency

Project Goals

- **Develop an integrated system classification of buildings engaged in the production of chemical and/or biological agents**
  - Physical, Biological and Computational
- **Hyperspectral and active imaging if bio-chem agents using “deep learning” algorithms**
- **Advance several remote spectroscopic technologies to the point of limited field deployment for signature identification**
- **Develop “Genetically Optimized Observant Bacteria for Amplifying Local Signatures” - (GOOBALS)**
- **Provide enhanced intelligence to the decision maker**

Project Goals (cont.)

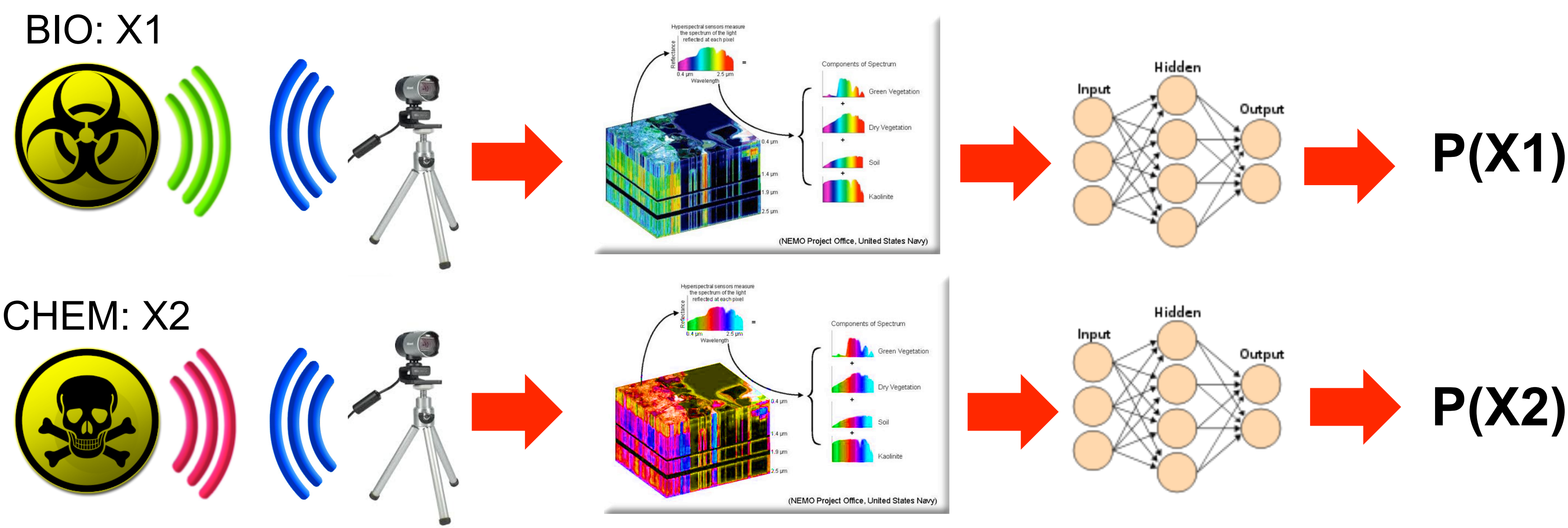




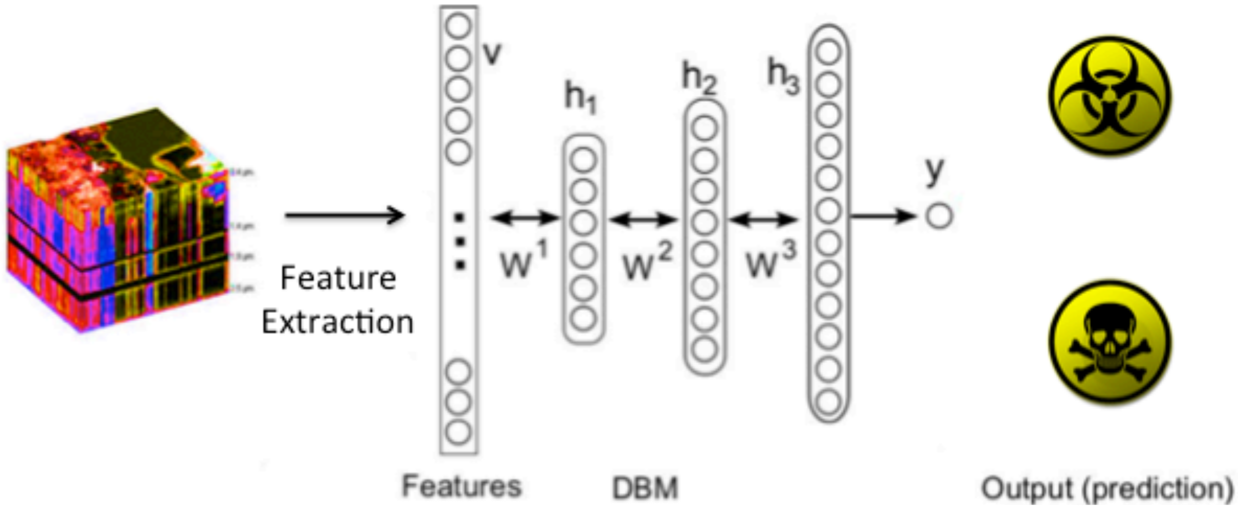
Laboratory Scale: Machine Learning

"Machine Learning, a branch of artificial intelligence, is about the construction and study of systems that can learn from data"
- wikipedia

a) Hazard b) Active/Passive Sensing c) Hypercube "Signatures" d) Machine Learning e) Classification



Laboratory Scale: Machine Learning



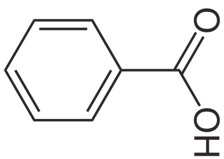
$$p(v,h)=\frac{1}{Z}\exp\left(v^T W^1 h^1 + h^{1T} W^2 h^2 + h^{2T} W^3 h^3 \right)$$

v = visible layer (the observed features)
h = "hidden" layer (features that capture the interdependencies of the observed measurements)
W = weights (connecting variables at adjacent layers)
Z = normalizing constant

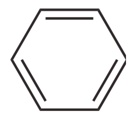
Laboratory Scale: Hyperspectral Training Data

- **Differentiation between chemicals of interest (“specificity”) can be difficult with a single probe**

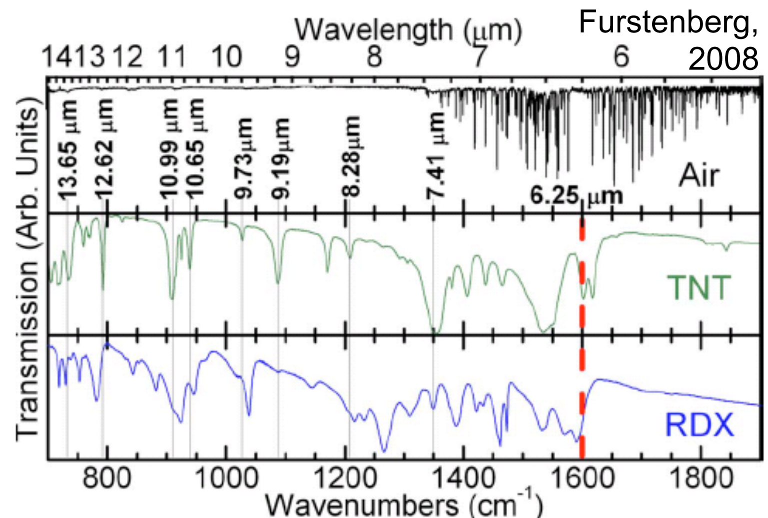
Food preservative

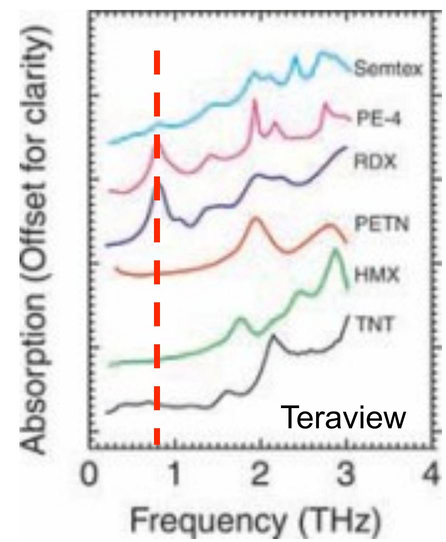


Toxic chemical



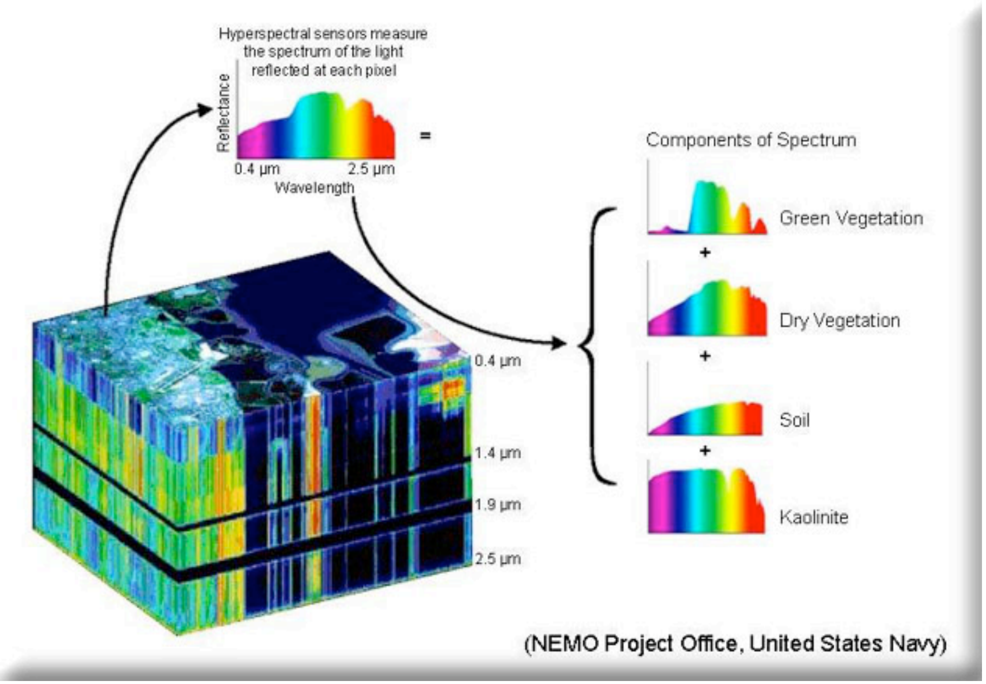
- **Current techniques have various advantages and disadvantages for forward deployment**
- **Combining multiple techniques can yield higher specificity and operational flexibility**
- **Visible, THz and IR can combine to specify TNT from RDX and PE-4**



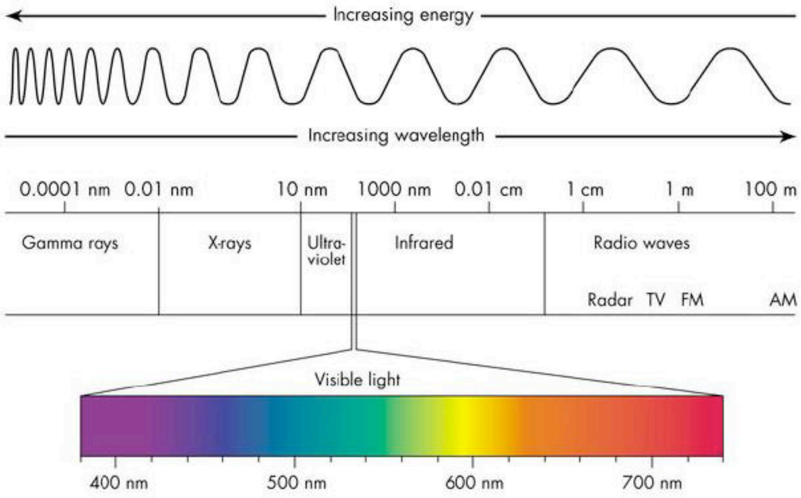


Laboratory Scale: Hyperspectral Training Data

- **Vibrational spectroscopic data from multiple regions of the electromagnetic spectrum can form a hypercube**

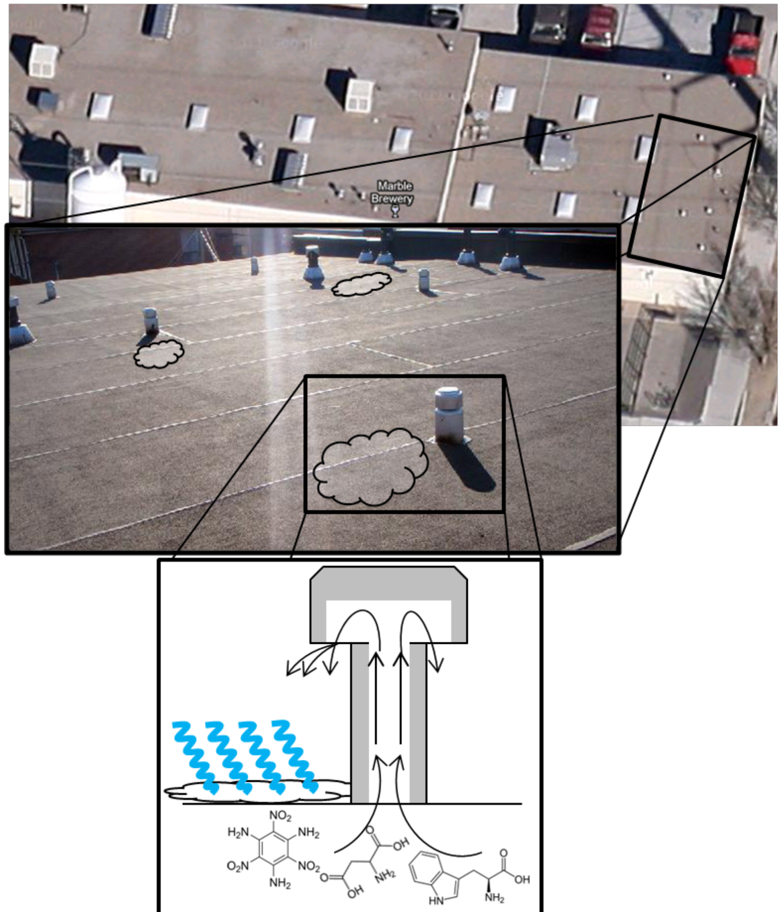


(NEMO Project Office, United States Navy)

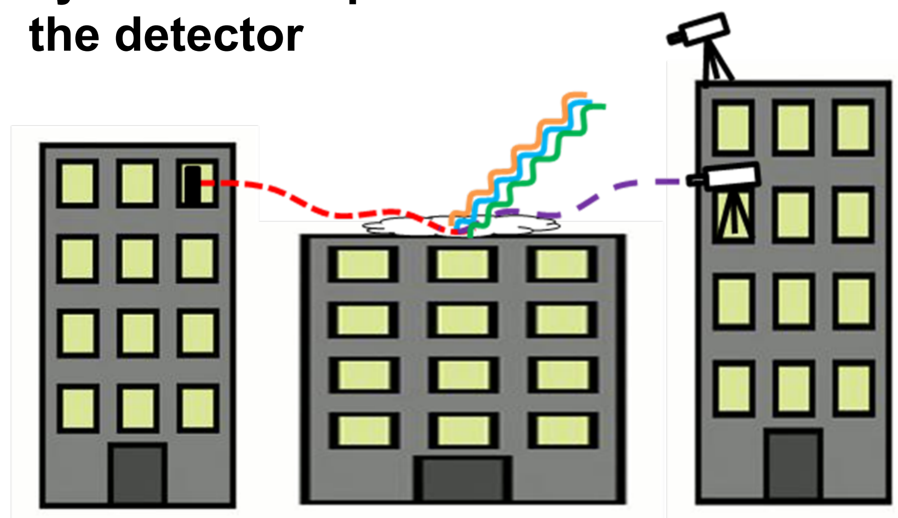


- **Even if there are no "100% certain signatures" in a given band, the combination of multiple bands will identify probable agents**

Deployment of Detection Technologies

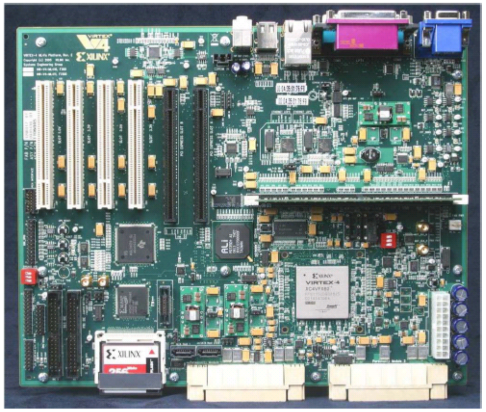


- Vibrational spectroscopy is the most mature field for remote or standoff detection of chemicals
- Current obstacles are generally engineering rather than science
- We will modify existing laser-based systems to separate the source from the detector

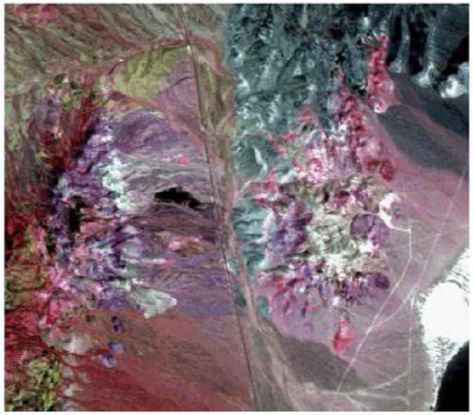


Deployment of Detection Technologies

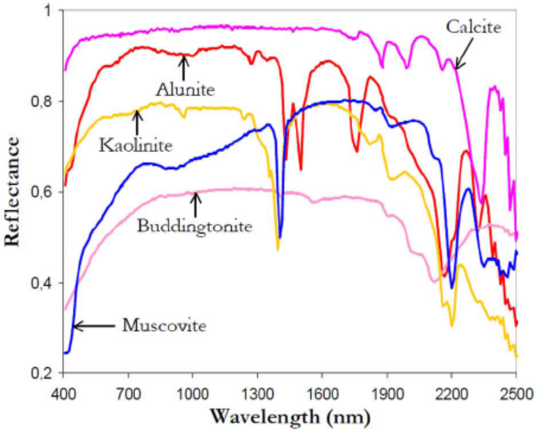
- **Field programmable gate arrays (FPGA) will couple the developed machine learning algorithms to hyperspectral detectors**



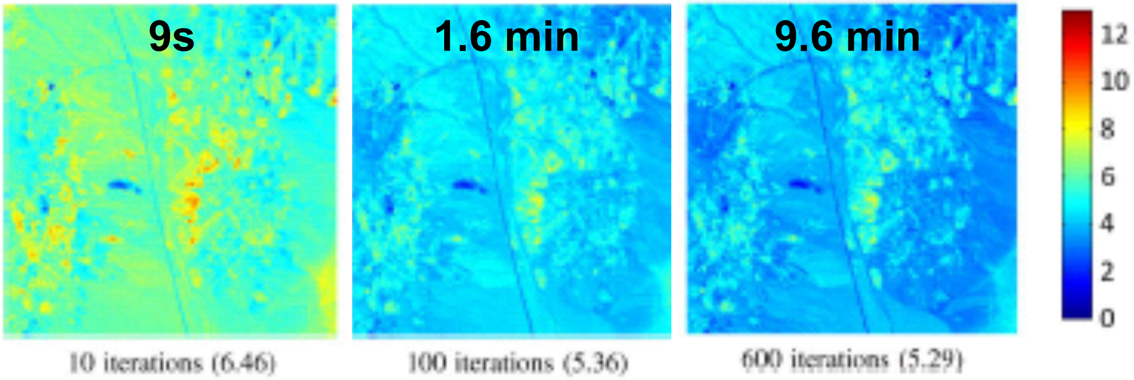
FPGA Board



Hyperspectral FC Image



Known Constituents



Gonzalez, 2012

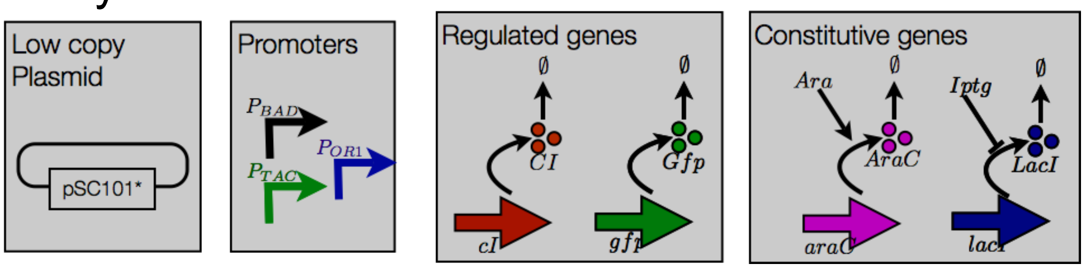
Design of Next Generation BioSensors

Passive signatures provide some useful information, but are often too non-specific and/or too weak to make critical decisions.

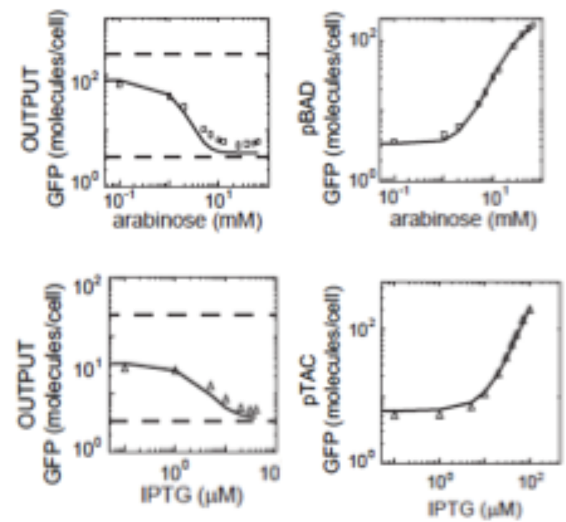
Active sensors can increase specificity and amplify weak signals.

Key Innovative Goal: **Use natural & synthetic organisms to amplify and communicate trace signals.**

Preliminary Work:



Computational/experimental design of synthetic genetic regulatory modules (above) lead to predictable input/output relationships (right).



Lou, et al, Nature Biotechnology, 2013

Design of Next Generation BioSensors

Subtasks:

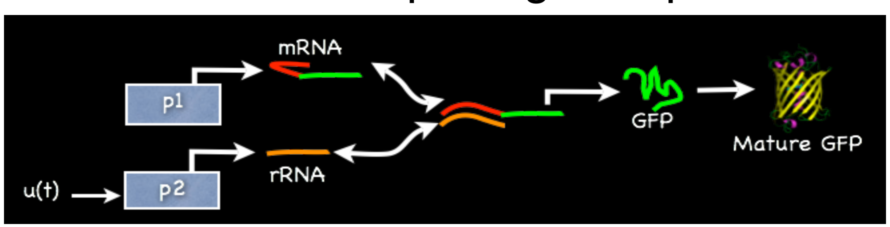
Mutate, screen and select RNA and protein aptamers to bind a library of orthogonal small molecules and biomarkers.

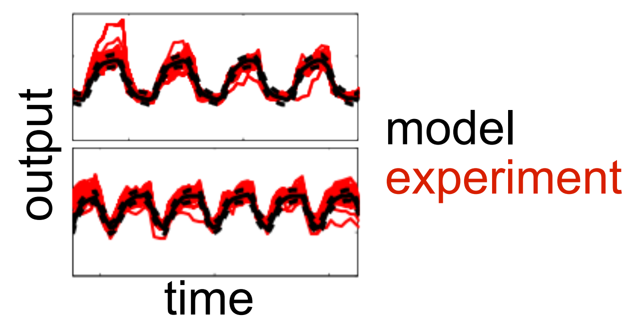
Screen and select for RNA and protein output products for standoff detection.

Optimize regulatory modules to process inputs and control cell response outputs.

Develop novel cell media to (1) promote synthetic cell growth for a short time, (2) amplify specific bacterial signals, and (3) enable remote deployment and rapid decay.

Integrate several rounds of modeling and experiment to predict local and standoff signatures of various input/logic/output/media combinations.





Summary

Machine Learning:
Innovation: Application of deep learning methodology to a high dimensional spectroscopic dataset
Expected Result: An efficient, predictive algorithm for substance classification

Hyperspectral Imaging:
Innovation: Integration of minaturized hyperspectral detectors with FPGA based on machine learning
Expected Result: Standoff detector capable of analyzing a target area (e.g. rooftop) in several minutes

Sensor Design:
Innovation: Use synthetic biology to sense/communicate biomarker presence.
Expected Results: Predictive understanding of local and standoff signatures of various combinations of input, logic, output and media.

Project Plan

- **Parallel efforts will proceed in hyperspectral detector optimization, synthetic biosensor design, and machine learning algorithm development.**
- **Efforts are independent yet additive—success in one aim will complement others, but failure in one will not hinder forward progress.**
- **We will systematically increase the scenario complexity**
  - 1) show that we can detect and identify some residue from a discarded roof tile
  - 2) show that we can do this from "far" away (e.g. 50 m)
  - 3) show that we can improve signature through active biosensor deployment
  - 4) employ machine learning algorithms to analyze remote hyperspectral data
- **Prototype sensors with a machine learning analysis package will be delivered by Year 3**
- **Transition to programmatic funding near the end of Year 3**

Team and Budget

- **Hyperspectral Imaging**
  - ISR, P23, …
- **Machine Learning**
  - ISR, CCS, CML
- **Advanced Sensors**
  - B10, B11
- **External Consultants**
- **Review committee**

Program development needs to develop a competitive proposal

- **Hyperspectral imaging - machine learning experiment (3 months)**
  - Two different point sources hypercubes (e.g. diesel vs. regular gas engines)
  - Training and validation of deep boltzman machine
  - Analysis and feedback
- **Feasibility studies of GOOBALS using existing organisms (6 months)**
  - Development of an adhesive media
  - Hyperspectral imaging of bacteria response to intended stimulus
  - Similar imaging of non-intended stimulus to establish probability of false positives
  - Field test of a simple scenario (e.g. deploy GOOBALS onto boxes containing chemicals)
- **Remote hypercube data collection using current technology (6 months)**
  - UAS-mounted detectors would image a simulated "suspicious facility"
  - Assess whether data from multiple EM bands is usable

Thank you!

Questions or Comments?